

## Article

---

« La puissance des images (note de recherche) »

Philippe Quéau

*Anthropologie et Sociétés*, vol. 16, n° 1, 1992, p. 91-94.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/015202ar>

DOI: 10.7202/015202ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

---

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <https://apropos.erudit.org/fr/usagers/politique-dutilisation/>

---

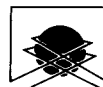
Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : [info@erudit.org](mailto:info@erudit.org)

# LA PUISSANCE DES IMAGES

(Note de recherche)

Philippe Quéau



On appelle image *numérique* une image qui se présente, à un moment de son processus de création, sous forme d'un tableau de nombres, rangés dans la mémoire de l'ordinateur. On distingue habituellement deux catégories d'images numériques : les images de « synthèse » et les images « traitées » par ordinateur. Les premières sont générées à partir de modèles logico-mathématiques et les secondes sont obtenues par des moyens de prise de vue classiques (dans le domaine de la lumière visible ou dans des gammes d'ondes invisibles comme les infrarouges ou les rayons X), qui sont ensuite numérisées afin de subir divers traitements mathématiques, comme des opérations de filtrage, d'extraction de caractéristiques, de reconnaissance de formes, etc.

Si les images photographiques ou vidéographiques restent issues du monde réel, c'est-à-dire de l'interaction de photons lumineux avec des surfaces photosensibles (pellicules photochimiques, tubes électroniques), en revanche les images de synthèse sont créées par des opérations abstraites, par la manipulation de langages symboliques ou logico-mathématiques. Ces manipulations symboliques obéissent à des règles formelles. On peut dès lors créer des images en formant des arrangements corrects de symboles, dans le cadre d'une syntaxe donnée. Un arrangement symbolique capable de générer une ou plusieurs images est appelé un « modèle ». Ce concept doit être compris dans le sens qu'il a lorsqu'on parle de modèle mathématique ou de modèle physique. Il ne s'agit pas nécessairement d'un modèle tentant de mimer quelque réalité préexistante. Il peut y avoir des modèles plus ou moins arbitraires et gratuits. Le plus souvent, les modèles sont conçus pour donner une forme mathématique à une idée, et pour permettre ainsi d'en simuler le fonctionnement, afin de la valider ou de la réfuter. Les modèles que l'on crée abstraitement ne trouvent de justification que dans leur propre cohérence interne.

Avec les techniques de programmation objet et autres développements issus des recherches sur l'intelligence artificielle, on peut également parler de modèles « symboliques », consistant dans le couplage de « base de connaissances » et de « règles d'inférences ». Les images générées à l'aide de ces modèles ne doivent plus être considérées simplement pour ce qu'elles donnent à voir, c'est-à-dire comme des images de quelque chose dont elles seraient la copie. Elles doivent être interprétées comme des « phénomènes » permettant d'accéder à une certaine compréhension du modèle qui les engendre. Elles jouent le rôle de « fenêtre » permettant de jeter un regard sur la réalité conceptuelle recélée par le modèle.

Il faut distinguer clairement ces deux mondes, celui des modèles et celui des images, le domaine de l'intelligence et celui du sensible. Synthétiser une image

consiste essentiellement à passer d'une représentation conceptuelle, formalisée par un modèle « intelligible », à une représentation concrète, sensible, l'image « visible ». En revanche, traiter une image, c'est effectuer la démarche inverse. À partir d'images déjà existantes, il s'agit d'extraire des éléments caractéristiques ou des régions pertinentes de l'image et de les analyser de façon à acquérir une meilleure compréhension, une meilleure intelligibilité *a posteriori* de l'image. La synthèse d'image concerne des applications de simulation, de conception assistée par ordinateur, de création graphique. Le traitement d'image intéresse les spécialités disposant d'un grand nombre d'images à analyser (astrophysique, médecine, télé-détection, etc.). Cependant, depuis quelque temps on voit apparaître des applications tendant à utiliser conjointement les deux voies : (modèle  $\rightarrow$  image) et (image  $\rightarrow$  modèle), comme dans certaines applications de vision par ordinateur, de reconnaissance des formes ou de poursuites de cibles.

Ceci nous amène à considérer les images numériques de trois façons différentes et complémentaires :

- comme des « images », dotées des mêmes contraintes de cohérence physique et symbolique que dans le cas des images classiques, et nous donnant un point de vue sur un ensemble d'objets visibles possédant des relations spatio-temporelles ;
- comme des « moments » dans l'histoire d'un processus récurrent, celui de la construction interactive de la relation même du modèle et des images qu'il engendre ;
- comme des éléments d'appréhension « sensible » de la forme « intelligible » qu'est le modèle.

Suivant les cas, on abordera l'image pour elle-même, ou bien pour sa capacité à constituer ou à faire évoluer le modèle ou enfin comme un moyen tangible de comprendre ce dernier.

Il importe de discriminer ces différents statuts de l'image, et surtout la façon dont ils influent sur la constitution ou la compréhension du modèle. En effet, les images numériques et les modèles logico-mathématiques peuvent être considérés comme des arrangements symboliques disposés dans des mémoires d'ordinateur. Ceci implique la possibilité concrète de passer, par de simples manipulations de symboles, du niveau de représentation du modèle au niveau de représentation de l'image, et réciproquement. Par exemple, un modèle disposant de critères de traitement d'une image peut générer des « prédictions » sur la présence ou l'absence de caractéristiques spécifiques dans l'image. Ces prédictions subissent alors une vérification qui donne lieu à une validation ou à une réfutation, à partir de l'image elle-même. Suivant les cas, une correction appropriée du modèle peut être envisagée, conduisant à une étape ultérieure du processus.

On peut donner l'exemple du système PRIAM développé par l'ETCA et appliqué à la poursuite de cible à l'aide de caméras infrarouges. On utilise la méthode d'analyse suivante : on sélectionne un petit nombre de points caractéristiques répartis autour des régions « chaudes » de l'image infrarouge. Ensuite, on construit des « régions » à partir de ces points jugés intéressants par croissance concentrique, en examinant si les points de ces voisinages font partie ou non de la région « chaude ». Ensuite, on cherche à regrouper les régions obtenues de façon à

en faire des « objets » appartenant au même ensemble de régions adjacentes. Ce sont les « parties d'objets », supposées être plus robustes que les points isolés, qui serviront de marqueurs de référence pour le traitement des images suivantes.

PRIAM effectue ainsi progressivement son « apprentissage », à partir de connaissances *a priori* sur les « cibles » censées être rencontrées dans l'image. Il sélectionne les « parties d'objets » ressemblant le plus au modèle idéal de cible qu'il possède en mémoire. Un module de « poursuite » effectue ensuite la mise en correspondance entre les points caractéristiques théoriques (remis à jour à chaque image) et les points effectivement observés. On utilise un algorithme de type prédiction/vérification. La prédiction consiste à générer des hypothèses d'association entre les « points » et les « régions ». La vérification consiste à calculer le « coût » de cette association. Ce coût est la somme d'un coût local (« ressemblance » entre deux points associés) et d'un coût global (« similarité géométrique » de deux ensembles de points). On peut alors déterminer une « zone de confiance » dans les régions de coût minimal. Ainsi, à partir de différentes sources d'images (radar, infrarouge, radio), on peut détecter des zones d'intérêt, mettre en valeur des indices visuels, extraire des contours et des arêtes, et donc « construire une représentation ». Par ailleurs, à partir d'un modèle traditionnel de la scène attendue et d'une analyse des premiers indices relevés, on peut choisir une « représentation de référence ». On cherche alors à mettre en correspondance la représentation « construite » et la représentation de « référence », de façon à conclure à une validation éventuelle du modèle, ou bien à induire d'autres hypothèses de représentation.

Ce jeu de prédictions et de vérifications successives entre la description *a priori* du modèle et la réalité effectivement observée permet d'améliorer la correspondance entre le modèle et l'image. Il y a « apprentissage » du système, ce qui permet de remplacer un premier modèle d'analyse de la scène par d'autres plus adéquats, en fonction des progrès réalisés dans les prédictions. On peut passer d'un modèle possédant des critères physiques d'analyse à un modèle aux critères plus cognitifs. Il s'agit d'une « commutation » entre modèles privilégiant des « visions du monde » différentes. On améliore progressivement la représentation de la scène, par une sorte d'épigénèse.

Dans le procédé de poursuite à partir de capteurs infrarouges, le modèle est d'abord physique et tient compte des modèles de propagation de la chaleur puisqu'il s'agit d'analyser des images infrarouges. Ceci permet une segmentation en régions, elle-même corollaire d'une détection de points jugés intéressants. Réciproquement, les régions extraites valident les points d'intérêt et ce sont ces derniers qui serviront à alimenter un deuxième modèle, plus conceptuel, susceptible de prédire le comportement des points caractéristiques des cibles poursuivies, en fonction d'une connaissance préalable de leur comportement.

L'exemple du système PRIAM fait référence à des techniques de traitement dont on pourrait penser qu'elles n'auraient pas d'utilité à s'appliquer aux images de synthèse. En effet, la plupart des modèles utilisés en synthèse d'image sont déterministes. Il s'agit de créer des images répondant à des fins précises (conception, création, animation...). Dans ce cas, la relation entre le modèle et l'image est parfaitement univoque et ne justifie en aucune manière l'application de

techniques de traitement ou d'analyse d'images. Cependant, certains modèles dits « semi-calculables » sont indéterminables. C'est le cas des automates cellulaires de classe 4 dans la classification de Stephen Wolfram<sup>1</sup>. Par exemple, les modèles d'automates asynchrones utilisent des règles d'itération pouvant elles-mêmes varier en fonction de l'état atteint par l'automate. Ces modèles sont théoriquement déterministes mais pratiquement indéterminables. Il faudrait en effet disposer d'un temps infini sur une infinité de machines de simulation pour explorer l'espace des phases d'un automate asynchrone. Il n'est donc pas possible de prédire le comportement exact de ces modèles ni même l'allure générale de ce comportement. Cependant, il reste intéressant de chercher des stratégies d'exploration des espaces des phases, afin d'en extraire un peu d'intelligibilité.

En soumettant les états successifs d'un automate à un modèle de traitement chargé de reconnaître des régularités ou des caractéristiques éventuelles, on cherche à utiliser l'étape actuelle de l'automate pour influencer les étapes futures. Les images produites par l'automate serviront à modifier les règles d'itération de ce dernier. Bien entendu, puisque ces images sont synthétisées à partir d'un modèle, on dispose de toute l'information « analytique » les concernant. Cependant, cette information ne porte que sur les états de chaque cellule et non pas sur l'ensemble des relations spatiales et des relations de voisinage qui relient toutes ces cellules entre elles. L'automate cellulaire résulte d'un ensemble de processus locaux interagissant localement et ne « sait » donc pas le rôle que chacun des processus générés joue par rapport à l'image globale qu'ils produisent. Le module de traitement d'image simule un « regard » sur l'image. On reconnaît des formes, on extrait des caractéristiques qui serviront à valider le fonctionnement de l'automate, ou alors à modifier ses règles d'itération pour en continuer l'exploration générale. Le traitement de l'image permet de valider le processus de sa synthèse.

En conclusion, l'hybridation aujourd'hui croissante des techniques de synthèse et de traitement de l'image numérique permet de créer des modèles évolutifs capables d'auto-apprentissage. Les images ne sont plus simplement cantonnées dans le domaine du paraître et de l'apparaître, dans la copie ou le simulacre, elles deviennent les actrices de leur mise en représentation. La puissance des images numériques tient à la béance signifiante, productrice entre leur formalisation intelligible et leur représentation sensible, béance sans cesse suturée par de constantes tentatives de réinterprétation.

Philippe Quéau  
Institut national de l'audiovisuel  
4, avenue de l'Europe  
94366 Bry-sur-Marne Cedex  
France

---

1. Voir notre analyse de ces automates dans *Metaxu. Théorie de l'art intermédiaire*, Seyssel, Champ Vallon, 1989.